

1 **Additional file 1**

2

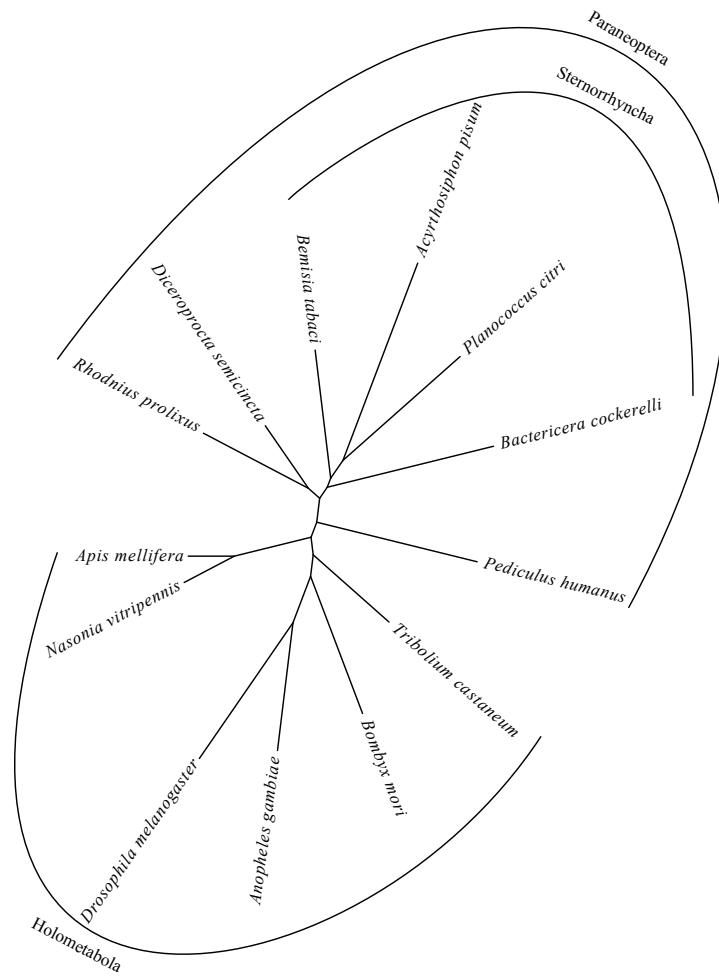
3 *Tree dating calibration and Bayesian priors*

4 The nodes used to calibrate the ML phylogeny (fig. S1) are identified in fig. S2 and the  
5 parameters used in the calibration are specified in table S2 [1].

6 *CAFE model fitting under alternative topology*

7 Models of gene family evolution were fitted in CAFE using a phylogeny with a  
8 paraphyletic Paraneoptera, similar to the topology in [1]. The best models were selected  
9 using likelihood-ratio tests, fitting the null distributions to  $\chi^2$  distributions with degrees of  
10 freedom equal to the difference in number of parameters between the two models.

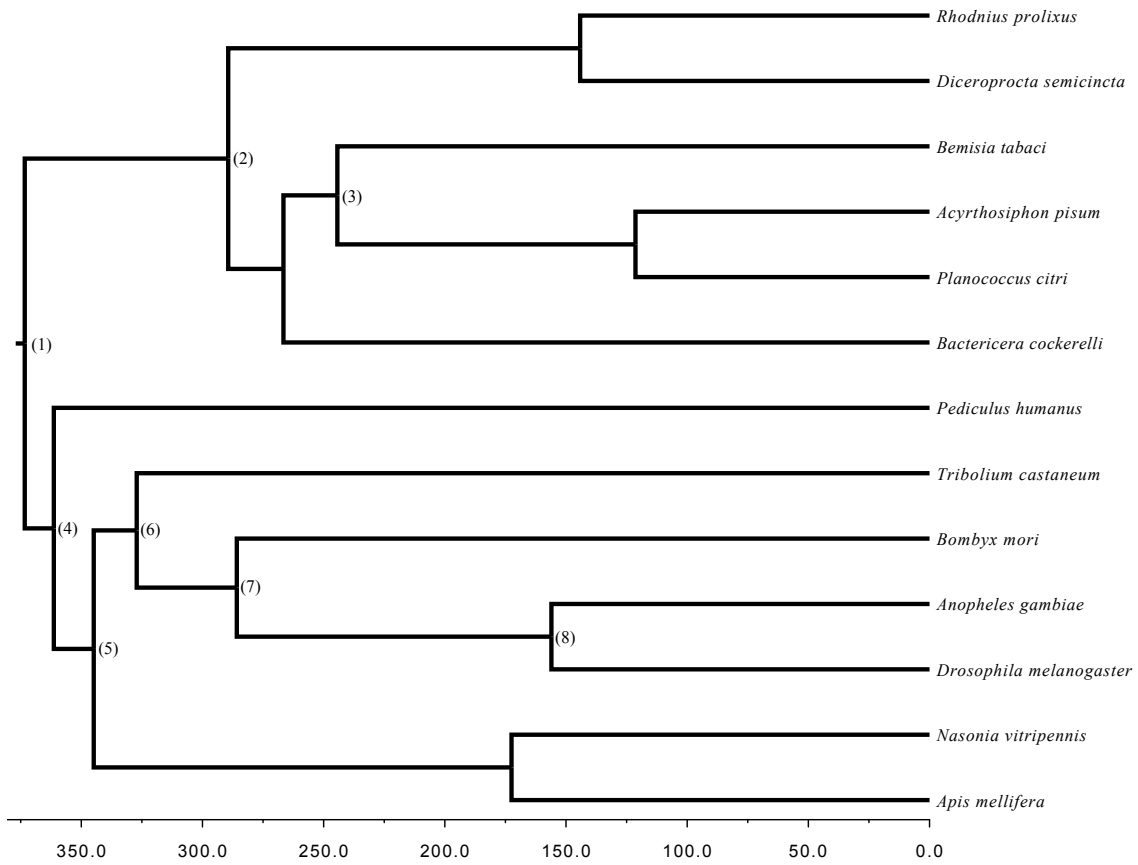
11



12  
13

Figure S1: Unrooted ML phylogeny inferred in RAxML.

0.2



14

15 Fig. S2: Calibrated phylogeny used as starting point of Bayesian inference. The labeled

16 nodes refer to the calibration point in table S2 used for the time calibration.

17

## References

- 18 1. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J,  
19 Flouri T, Beutel RG, Niehuis O, Petersen M, Izquierdo-Carrasco F, Wappler T, Rust J,  
20 Aberer AJ, Aspöck U, Aspöck H, Bartel D, Blanke A, Berger S, Böhm A, Buckley TR,  
21 Calcott B, Chen J, Friedrich F, Fukui M, Fujita M, Greve C, Grobe P, et al.:  
22 **Phylogenomics resolves the timing and pattern of insect evolution.** *Science* 2014,  
23 **346:763–768.**

24

25 Table S1: Taxonomy of the species used in the phylogenetic inference.

Species	Superfamily	Sub-order	Order	Super order
<i>Acyrtosiphon pisum</i>	Aphidoidea			
<i>Planococcus citri</i>	Coccoidea			
<i>Bemisia tabaci</i>	Aleyrodidea	Sternorrhyncha		
<i>Bactericera cockerelli</i>	Psylloidea		Hemiptera	Paraneoptera
<i>Rhodnius prolixus</i>	Reduvidoidea	Heteroptera		
<i>Diceroprocta semicincta</i>	Cicadoidea	Auchenorrhyncha		
<i>Pediculus humanus</i>		Anoplura	Phthiraptera	
<i>Nasonia vitripennis</i>	Chalcidoidea			
<i>Apis mellifera</i>	Apoidea	Apocrita	Hymenoptera	
<i>Tribolium castaneum</i>	Tenebrionoidea	Polyphaga	Coleoptera	
<i>Bombyx mori</i>	Bombycoidea	Glossata	Lepidoptera	Holometabola
<i>Drosophila melanogaster</i>	Ephydroidea	Brachycera		
<i>Anopheles gambiae</i>	Culicoidea	Nematocera	Diptera	

27 Table S2: Nodes and divergence times used to calibrate the ML phylogeny. Node number refers  
28 to the labels in fig. S2. All dates in million years before present.

Node	Divergence lower bound	Divergence upper bound	Calibrated value
1	346	401	373.33
2	233	346	289.38
3	182	307	244.34
4	334	390	361.32
5	318	372	344.87
6	301	353	327.13
7	244	329	285.83
8	107	207	156.48

29

30

31 Table S3: Akaike Information Criterion (AIC) of models in the DupliPHY-ML algorithm.

	Model	# Parameters	ML	AIC
APC	Parsimony	1	24.13	50.26
	Parsimony+G	2	24.32	52.65
	BDO	2	23.88	51.76
	BDI	3	23.68	53.37
	BDI+G	4	23.88	55.76
	BDIE	4	23.68	55.36
	BDIE+G	5	23.88	57.75
AAP	Parsimony	1	24.46	50.92
	Parsimony+G	2	24.66	53.31
	BDO	2	23.95	51.90
	BDI	3	23.75	53.50
	BDI+G	4	23.95	55.90
	BDIE	4	23.74	55.48
	BDIE+G	5	25.13	60.26

32

33

34

35

36

37

38

39 Table S4: Results of a Likelihood-ratio test comparing models with a global  $\lambda$  rate of gene duplication to models with a rate shift in  
 40 Sternorrhyncha. The models were fitted using a phylogeny with a paraphyletic Paraneoptera (fig. S2).  $\lambda$  represents the probability of  
 41 gene duplication/loss per gene per million years.

	Model	$\lambda_{\text{Background}}$ $\times 10^3$	$\lambda_{\text{Sternorrhyncha X}}$ $10^3$	Fold increase	-ML	# Parameters	Likelihood ratio	<i>p</i> -value
APC	Single $\lambda$	1.419	-	-	35.27	1	2.845	0.097
	Multiple $\lambda$	0.7235	3.4201	4.73	33.85	2		
AAAP	Single $\lambda$	1.3793	-	-	38.41	1	6.039	0.018
	Multiple $\lambda$	0.3774	3.7511	9.94	35.39	2		



43 Table S5: Results of the Likelihood-ratio test comparing models differentiating  $\lambda$  and  $\mu$  to models with equal birth-death parameters.  
 44 For APC, the estimated parameters are global across the phylogeny. For AAAP, the parameters were rates were allowed to shift in  
 45 sternorrhynchan species. The models were fitted using a phylogeny with a paraphyletic Paraneoptera (fig. S2).  $\lambda$  represents the  
 46 probability of gene duplication/loss (Single/Multiple  $\lambda$ ), or gene duplication only (Single/Multiple  $\lambda + \mu$ ) per gene per million years.  $\mu$   
 47 represents the probability of gene loss per gene per million years.

	Model	$\lambda_{\text{Background}} \times 10^3$	$\lambda_{\text{Sternorrhyncha}} \times 10^3$	$\mu_{\text{Background}} \times 10^3$	$\mu_{\text{Sternorrhyncha}} \times 10^3$	-ML	# parameters	Likelihood ratio	<i>p</i> -value
APC	Single $\lambda$	1.419	-	-	-	35.27	1	0.03	0.855
	Single $\lambda + \mu$	1.3923	-	1.4517	-	35.25	2		
AAAP	Multiple $\lambda$	0.3774	3.7511	-	-	35.39	2	8.57	0.011
	Multiple $\lambda + \mu$	$1.14 \times 10^{-10}$	3.01	0.5096	1.8965	31.11	4		